I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10991

Activation against Absenteeism: Evidence from a Sickness Insurance Reform in Norway

Øystein Hernæs

SEPTEMBER 2017

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10991

Activation against Absenteeism: Evidence from a Sickness Insurance Reform in Norway

Øystein Hernæs
Institute for Social Research and IZA

SEPTEMBER 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Activation against Absenteeism: Evidence from a Sickness Insurance Reform in Norway*

I evaluate a program aimed at strictly enforcing a requirement that people on long-term sick leave be partly back at work unless explicitly defined as an exception. Employing the synthetic control method, I find that the reform reduced work-hours lost due to absenteeism by 12 % in the reform region compared to a comparison unit created by a weighted average of similar regions. The effect is driven by both increased part-time presence of temporary disabled workers and accelerated recovery. Musculoskeletal disorders was the diagnosis group declining the most. The findings imply large savings in social security expenditures.

JEL Classification: I18, I38, J48

Keywords: absenteeism, disability, activation, forkfare

Corresponding author:
Øystein Hernæs
Institute for Social Research
Postboks 3233 Elisenberg
0208 Oslo
Norway
E-mail: ohernaes@gmail.com

* Funding from the Røwde foundation and the project "Explaining social patterns in sicknes absence: the influence of values, attitudes and norms" (Norwegian Research Council grant #237993) is gratefully acknowledged. I wish to thank Tarjei Havnes and Knut Røed for valuable comments at earlier stages of the project.

NON-TECHNICAL SUMMARY

Rising disability rolls

Disability rolls have been rising for decades in many OECD countries, entailing both a substantial volume of labor withdrawn from the market, as well as heavy social security costs. This has led to increased attention on the trade-off between generosity towards those hit by a negative health shock and potential moral hazard problems that faces any social security system. Traditional responses to this trade-off have been to establish strong screening criteria or other gatekeeping policies, or to limit the level or duration of benefits.

Activation reform

This paper tries to answer whether an activation strategy based on graded sickness insurance, i.e. requiring temporary disabled workers to be partly back at work to the extent possible, as opposed to any absence automatically being 100%, can help reduce absenteeism and curb the corresponding social security costs. The paper analyzes a program implemented in the Norwegian region of Hedmark in 2013 aimed at strictly enforcing an already existing requirement that an employee on long-term sick leave be partly back at work unless explicitly judged by a physician to be unable to work at all, irrespective of adaptions at the workplace.

Results

The results show that the program to make use of the partial work capacity of workers on long-term sick-leave reduced absenteeism by 12 percent and brought large savings to the social insurance system. The effects were remarkably similar across gender and age groups. The estimated effect is stable across gender and age groups, and somewhat smaller in the construction sector. I find evidence that the absence rate declined not only through exploiting the partial work capacity of temporary disabled workers, but also by speeding up the transition rate back to full-time work. Consistent with expectations, the largest decline occurred for absenteeism due to musculoskeletal disorders, the smallest for respiratory disorders, with diagnoses for psychological and other disorders in between.

Policy implications

Such an activation strategy represents an alternative to traditional attempts at welfare reform involving stricter screening or reductions in generosity, and may be more compatible with already existing legislation and contractual obligations, as well as easier to find support for across political priorities.

1. Introduction

Can an activation policy targeted at making use of the partial work capacity of temporary disabled workers reduce long-term absenteeism and bring down social security costs?

Disability rolls have been rising for decades in many OECD countries, entailing both a substantial volume of labor withdrawn from the market, as well as heavy social security costs (OECD, 2010). This has lead to increased attention on the trade-off between generosity towards those hit by a negative health shock and potential moral hazard problems that faces any social security system. Traditional responses to this trade-off have been to establish strong screening criteria or other gatekeeping policies, or to limit the level or duration of benefits. In this paper I analyze whether an activation strategy based on graded sickness insurance, i.e. requiring temporary disabled workers to be partly back at work to the extent possible, as opposed to any absence automatically being 100%, can help reduce absenteeism and curb the corresponding social security costs.

Since 2004, the Norwegian social insurance system has required that an employee on long-term sick leave be partly back at work unless explicitly judged by a physician to be unable to work at all, irrespective of adaptions at the workplace. However, despite the clear letter of the law and pronounced political intentions, this rule has not been rigorously enforced. I analyze a program implemented in the Norwegian region of Hedmark in 2013 aimed at more strictly enforcing this already existing requirement.[1] There are clear indications that the reform was succesful in reducing absenteeism: a report from the Labour and Welfare Administration computed a fall of 8 % in Hedmark relative to the countrywide average in the year following the reform (Kann et al. 2014). In this paper, I conduct a rigorous analysis of a longer time period, consider the uncertainty and robustness of the results, and go deeper into mechanisms.

[1] Norway has 19 regions with an average population of 269,000. Hedmark has 194,000 inhabitants.

The fact that the program was comprehensively implemented across a whole region in a top-down fashion makes it natural to base an effect evaluation on some sort of comparison with other regions, thus the choice of comparison group becomes crucial. To make this choice, I employ the synthetic control method (SCM) due to Abadie and Gardeazabal (2003) and Abadie et al. (2010, 2015), which provides a data-driven approach to the selection of comparison group. The SCM matches the treatment region Hedmark and the control regions on several pre-treatment variables, in order to construct a counterfactual "synthetic" Hedmark based on a convex combination of regions, which is similar to the actual Hedmark, but that did not implement any changes to their activation policies.

The comparison of the actual region versus the constructed counterfactual shows that enforcing the graded sick-leave requirement reduced the sickness absence level by 12 %. Systematic placebo tests on all other regions and for an earlier time period find no estimated effects of comparable magnitude. The effect is also robust to successively leaving out the regions receiving a positive weight in the baseline analysis. The estimated effect is stable across gender and age groups, and somewhat smaller in the construction sector. I find evidence that the absence rate declined not only through exploiting the partial work capacity of temporary disabled workers, but also by speeding up the transition rate back to full-time work. Consistent with expectations, the largest decline occurred for absenteeism due to musculoskeletal disorders, the smallest for respiratory disorders, with diagnoses for psychological and other disorders in between. The findings imply savings in social security expenditures of USD 310 per employee.

2. Related literature

The paper is related to the literatures on activation strategies and return-to-work interventions for sick or temporary disabled workers, see Røed (2012) and van Vilsteren et al. (2015), respectively, for reviews. Graded sick leave has been growing in popularity as a policy tool in reccent years, and its impact has been investigated in several Northern European countries. Høgelund et al. (2010) and Schneider et al. (2016) find that being on graded sick leave is associated with a substantially higher chance of returning to regular working hours in Denmark and Germany, respectively. Likewise, Andrén and Svensson (2012) and Andrén (2014) show positive associations between graded and return to full-time work also for a sample of Swedish sick-listed employees. Analyzing a small (N=63) randomized controlled trial in Finland, Viikari-Juntura et al. (2012) find that grading accelerated the return to regular work of people suffering from musculoskeletal disorders. Also in Finland, Kausto et al. (2012) and Kausto et al. (2014) use propensity score matching to compare people on partial vs. full-time sick leave, and find that partial sick leave is associated with a larger share of people on partial rather than full-time disability pension and a higher degree of employment. A limitation of these studies is that they rely on relatively strong modeling assumptions and/or small samples.

A study in a different modeling tradition is Markussen et al. (2012), who also study activation requirements for workers on long-term absence in Norway. Using administrative register data from the first half of the 2000's, they employ an instrumental variable approach based on variation in physicians' propensity to use graded absence certificates in contrast to awarding full-time absence certificates. Their results are encouraging – requiring workers on long-term sick leave to work part-time decreases absenteeism and raises later employment propensities. These results are highly policy-relevant, as they are well identified and based on a scalable treatment that consists of a broad activation policy not restricted to certain therapies or diagnoses.

This paper complements Markussen et al. (2012) in three important respects: First, they estimate a local average treatment effect based on patients whose grading outcomes are influenced by the physician. Second, their main analysis is conducted in a context where graded absence certificates were relatively rare. Indeed, when they restrict their sample to after July 2004, when a reform sharply increased the use of graded absence certificates, the estimated effects are smaller, raising the worry of diminishing returns to the favorable effects of grading. Third, although the results are stable across a range of robustness checks, the chance remains of positive selection of patients to physicians with higher grading propensities, or that unobserved physician factors might influence both grading strategies and patient outcomes. It is therefore reassuring that this paper, based on a completely different identification strategy, in a setting in which graded absence certificates were much more common, also find substantial beneficial effects.

3. Activation Reform

In the Norwegian sickness insurance system, the replacement ratio is 100% from the first day of sick leave and up to one year. The employer pays for the first 16 days, while the state covers the period thereafter. For absence spells of more than 4 days, a physician must certify in writing that the employee has a medical condition that disables work. People on long-term sick leave (more than 8 weeks) are disabled in the sense that they have some condition that disables them from doing their job in a regular way. However, their situation is temporary, as the maximum long-term sick leave is 52 weeks, after which more permanent disability or retraining options are investigated.

As a measure to limit moral hazard problems on part of both the employer and the employee as well as to accelerate recovery (Mykletun et al., 2010), since 2004 there has been a *de jure* activation requirement for people on sick leave: They are required to be partly back at work as early as possible and at the latest after 8 weeks. This is based on the fact that most long-term sick-leave days are caused by musculoskeletal diseases, back pain, mental disorders or other illnesses that rarely are completely disabling and do not risk infecting other workers (Markussen et al., 2012), and for which some degree of activity may even be healthy (Hagen et al., 2003; Waddel and Burton, 2006). The activation requirement is aimed at exploiting the remaining work-capacity of the temporary disabled worker. When writing an absence certificate, which is necessary to receive sick-pay, physicians are obliged to report the fraction of the patient's work capacity that is lost due to the illness. A graded certificate indicates that this fraction is less than 100%. In the case of 30% lost work capacity, the worker would be obliged to perform 70% of her/his regular work duties, for which regular wages would be paid, and receive sick pay for the remaining 30%.

After 8 weeks of sick leave, the local social security administration is supposed to decide formally on whether a case is an exception to the activation requirement. An exception due to medical reasons should be based on an explanation by the employee's physician for why a graded certificate cannot be used. Likewise, an exception due to the impossibility of implementing sufficient adaptions at the workplace should rest on an explanation by the employer for why workplace adaptions are not possible. In the vast majority of exceptions in 2012, medical reasons were given as the basis of the exception (Brage et al., 2014).

However, the requirement that an absentee be partly back at work – activation – has *de facto* been far from rigorously enforced – the social security administration often did not formally decide on exceptions at the 8-week point, and accepted incomplete explanations for exceptions from physicians and employers. On the basis of this knowledge, the social security administration in the region of Hedmark set out to enforce the activation requirement comprehensively and consistently throughout the region starting from the second quarter of 2013.[2]

The implementation consisted of three main parts: First, establishing a system for supporting individual case-workers regarding the new practice, including regular meetings with the purpose of discussing all relevant cases. In these meetings, case-workers prepare their cases and discuss them with administrative and medical support staff from the social security administration. Second, consistently sending a letter to the (full-time) absentee around 8 weeks informing about termination of the sick leave benefit unless satisfactory reasons for an exception are given, as well as a change in the wording of the standard text. The rewriting consisted of encouraging contact with the employer rather than the physician. Third, information about the reform to the parties involved: Letter and e-mail to all physicians, information to employers from the region's "workplace support center", and information to the general population through local newspapers.

4. Method and data

The activation reform was comprehensively implemented across the whole region of Hedmark from the second quarter of 2013. From a policy evaluation perspective, the fact that the implementation took place at a specific point in time determined by the central authorities is helpful, as it allows a credible comparison of the pre- and post-periods. Yet, absenteeism levels move substantially over time also for other reasons, necessitating a comparison of the development from pre- to post-period with deveopments elsewhere. One solution would be a difference-in-differences (DiD) analysis, comparing changes in outcomes across time and regions. However, in this case there is only one treated region, which makes the regular re-

[2] The account of the reform relies on NAV Hedmark (2014).

gression framework problematic. Moreover, a simple average of other regions may not consitute the most relevant comparison group. Furthermore, though the DiD framework does the important job of differencing out individual-specific heterogeneity that is fixed in time, time-varying factors at the individual level may also be important.

To deal with these concerns, I employ the synthetic control method (SCM) due to Abadie and Gardeazabal (2003) and Abadie et al. (2010, 2015) to construct the counterfactual. The goal is to esimate the effect of the intervention/treatment for unit *i* at time *t*: $\alpha_{it} = Y_{it}^{I} - Y_{it}^{N}$, where Y^{I} denotes the outcome of interest in case of exposure to the intervention, and Y^{N} denotes the outcome in the absence of intervention. The SCM is motivated by a structural model in which the non-intervention outcome for a unit *i* is given by:

$$Y_{it}^{N} = \delta_t + \boldsymbol{\theta_t Z_i} + \boldsymbol{\lambda_t \mu_i} + \varepsilon_{it}, \tag{1}$$

where δ_t is a common time-dependent factor, $\boldsymbol{\theta_t}$ is a vector of unknown parameters, $\boldsymbol{Z_i}$ is a vector of observed covariates not affected by the intervention, $\boldsymbol{\lambda_t \mu_i}$ is a vector of unobserved time-specific common factors multiplied by a vector of unobserved unit-specific factor loadings, and ε_{it} is an unobserved transitory shock. Note that the term $\boldsymbol{\lambda_t \mu_i}$ allows the effect of unobserved unit-specific confounders to vary over time. Next define the unit of interest as unit *i*=1, and a vector of weights $\boldsymbol{W} = (w_2, \dots, w_{J+1})'$, such that $w_j \geq 0$ and $w_2 + \cdots + w_J = 1$, defined over all *J* non-treated units. Any **W** represents a possible synthetic control and can for any time period be used to produce a linear combination of non-treated units $\sum_{j=2}^{J+1} w_j Y_{jt}$ to serve as a counterfactual for the intervention unit. Abadie et al. (2010) show that in the context of the model specified in Equation (1), a **W*** such that $\sum_{j=2}^{J+1} w_j^* Z_j = Z_1$ and $\sum_{j=2}^{J+1} w_j^* \overline{Y}_J = \overline{Y}_1$, where $\overline{Y}_i$ is a linear combination of pre-intervention outcomes, would provide an unbiased estimator of Y_{1t}^{N}. This solution is approximated by minimizing the distance between these so-called predictor variables of the treated unit and the synthetic control unit with respect to **W*.**

In this procedure, the weights put on the predictor variables are chosen to minimize the mean square predictor error of the outcome variable for the pre-intervention (or pre-validation) periods.

The result is a counterfactual consisting of a convex combination of non-treated ("donor") regions. In this paper, I employ averages of the outcome variable and the workforce and age structure of the region as the predictor variables to be used in the construction of **W***. These averages are defined over the period ranging from the first quarter of 2008 until the fourth quarter of 2011. This is also the period for which the mean squared prediction error of the outcome is minimized. The four quarters of 2012 and the first quarter of 2013 serve as a validation period for the weights constructed from the earlier period, before the treatment is introduced in the second quarter of 2013.

The procedure desctibed above provides an estimate of the counterfactual development of the treatment region, and thus of the treatment effect for that region, however, it does not say anything about the uncertainty of that estimate. As the number of regions is relatively small, standard large-sample inference is not applicable. I follow Abadie et al. (2010), who based on the idea of a permutation test suggest systematic placebo tests on all untreated units to see whether the estimated effect for the treatment unit is large relative to the placebo effects for the untreated units. Following Abadie et al. (2015), I complement these placebo tests "in space" with a placebo test "in time," moving the reform forward to the middle of the pre-treatment period, as well as a robustness exercise consisting of leaving out the regions receiving positive weight in the baseline results.

All variables are region-level data aggregated from individual-level register data by Statistics Norway. The main outcome is the sickness absence rate, defined as "man-days lost due to own sickness as a percentage of contractual man-days." This rate expresses the amount of work hours lost to absenteeism based on measuring the "stock" of physician-certified sickness absence on specific reference days. I use data starting in 2008, as there was a break in the data series at that point regarding industry classification scheme. However, the reform that I study was not implemented until the second quarter of 2013, which leaves 21 quarterly pretreatment observations. I study the period up to and including the fourth quarter of 2015. In the main analysis I focus on all employees (i.e. everyone aged 16-69), but I also break the analysis down by gender, age, sector and diagnosis group. It should be noted that as the outcome or subgroup changes, the synthetic control will typically also change, thus these estimates should be interpreted with more caution. To provide some information about the uncertainty also of these secondary estimates, the estimated effects will be displayed visually together with the placebo estimates for the other units.

The covariates on which I match are the share of inhabitants with a university or college education, the share of the workforce employed in the health or social sector and the public sector the female share of the workforce, and the age composition of the inhabitants.

There are 19 regions in Norway. The regions are the second administrative level of the country, below the national level. The main responsibilities of the region governments are upper secondary education, regional roads, local public transportation and some cultural and health services (Fiva and Halse, 2016). The Norwegian Labour and Welfare Administration has offices in all 428 municipalities of the country (the third administrative level), however these are internally organized under a region-level leadership, which may exert considerable influence on practice in the municipality offices. A map of the regions is provided in Figure A.1 in the appendix. I am forced to exclude the region of Oppland, as it introduced a similar activation program some time later, thus I am left with 18 regions - the treatment region of Hedmark, and the remaining 17 regions constituting the donor pool from which a counterfactual will be created.

5. Results

5.1. Main results

Figure 1 shows the time series for Hedmark and its synthetic counterpart. The synthetic control region closely traces the actual one for the whole pre-treatment period, i.e. until and including the final quarter of 2011, even including the various seasonal cycles. The fit from the out-of-sample prediction provided in the validation period, i.e. all quarters of 2012 as well as the first quarter of 2013, is also quite tight, before the two graphs sharply diverge at the introduction of the reform.

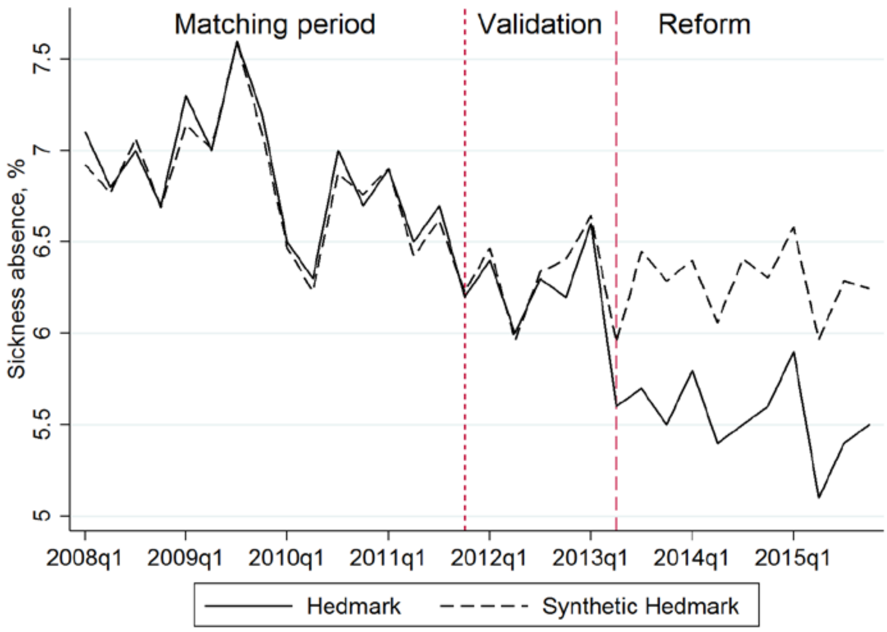

Figure 1. Trends in the sickness absenteeism in Hedmark and the synthetic control region.
Note: The dotted line at the fourth quarter of 2012 indicates the final quarter of the matching period. The dashed line at the second quarter of 2013 indicates the period in which the activation program was introduced.

Plotting the difference between the two graphs in Figure 2, we again see the close match until the first period of 2013, before a substantial difference sets in from exactly the quarter in which the activation program was introduced, drops further to around –0.75 percentage points in the following period, and remains at roughly that level for the rest of the time period. The difference for the whole post-treatment period is –0.72 percentage points, or –12 % compared to the counterfactual post-treatment level, ref. Table 1. One common worry with graded sick leave is that it may bring down absenteeism in the short run, but increase it in the long run as people are not given the appropriate rest to recover. It is therefore reassuring that there is no sign that the effect is reversed or even declines with time.

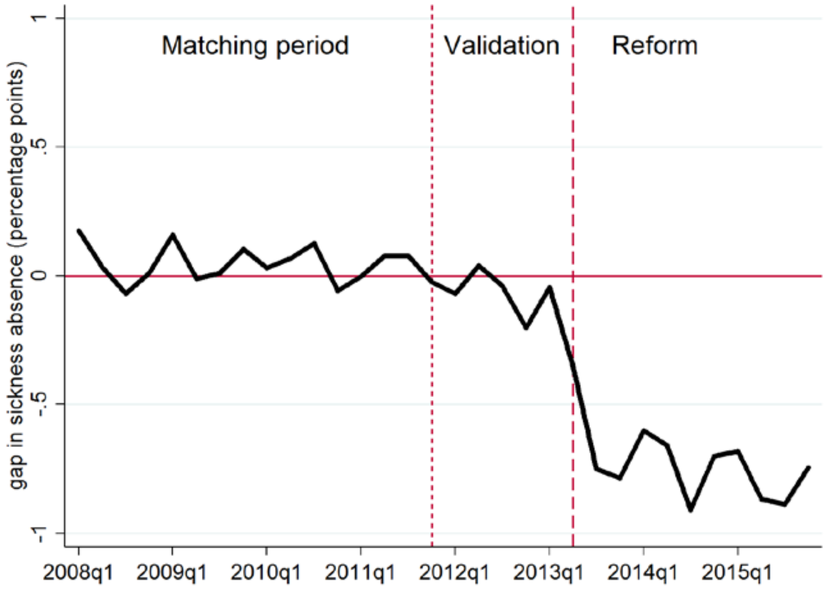

Figure 2. Difference in sickness absenteeism between Hedmark and the synthetic control region.
Note: The dotted line at the fourth quarter of 2012 indicates the final quarter of the matching period. The dashed line at the second quarter of 2013 indicates the period in which the activation program was introduced.

Estimated difference in post-treatment sickness absenteeism	
Synthetic Hedmark	6.267
Hedmark	5.545
Difference, absolute	–0.722
Difference, %	–12

Table 1. Difference in sickness absenteeism in Hedmark and the synthetic control region after treatment.

The first columns of Table 2 shows the values of the predictor variables for the treatment region – the real and synthetic Hedmark. The match is very close. To investigate whether there is a danger of interpolation bias, the final three columns list the values of the predictors for all the three regions receiving positive weights from the matching procedure (Nordland, Østfold, and Telemark). The weights are shown in the column headings. Nordland receives the largest weigh – 56.8 %, and is particularly close to Hedmark on all variables. Also the other two regions are quite similar, thus interpolating from very different regions is not a problem.

	Treatment region		Donors		
	Hedmark		Nordland	Østfold	Telemark
Weight	Real	Synthetic	0.568	0.353	0.079
Sickness absence, %	6.844	6.800	6.694	7.063	6.388
Share of workforce					
…with university education	0.256	0.257	0.257	0.256	0.259
…in health and social services	0.219	0.216	0.224	0.203	0.218
…in public sector	0.073	0.068	0.078	0.056	0.051
…females	0.526	0.528	0.529	0.528	0.524
Share of inhabitants					
…aged 20-29	0.104	0.112	0.113	0.110	0.113
…aged 30-39	0.120	0.125	0.119	0.133	0.126
…aged 40-49	0.142	0.144	0.143	0.147	0.142
…aged 50-59	0.140	0.134	0.136	0.132	0.137
…aged 60-69	0.122	0.113	0.113	0.113	0.115

Table 2. Means of predictor variables from the matching period.

To get a sense of the economic significance of the estimated effect, we can compute the number of full-time equivalent working days the reform saved. Contracted hours in Hedmark in the post-treatment period was in average 4 000 000 days of work per quarter, thus a reduction in absenteeism of 0.72 percentage points constitutes close to 30 000 working days saved per quarter, or 120 000 days per year. Considering the fact that the replacement rate in the Norwegian system is 100 percent (up to a ceiling), it is clear that there are large savings involved for the public purse. How large? 120 000 working days per year constitute around 520 man-years. The average sickness benefit basis in 2014 was USD 47,000 (NAV, 2016). Thus for one year, gross savings, in Hedmark only, are given by 520* USD 47,000 = 25 million, or USD 310 per employee. Since sick pay is taxed at the same rate as regular income, this number is also the net savings. To the extent that these results might indicate lower permanent disability benefit uptake later on, the savings could magnify considerably.

The costs are more difficult to estimate. Even though in many cases a partly available employee will be preferable to no employee at all, in general, adapting to a worker with a graded sick leave certificate does carry costs for an employer. There are adaption costs also for the employee, although for the employee it is more likely that the long-term benefits of keeping a relation to the employer, quicker recovery and avoiding earnings losses (Markussen, 2009) may outweigh the costs. Finally, the social security administration did incur some extra personnel costs related to the reform, but mostly handled the process by reallocating existing resources (NAV Hedmark, 2014). Thus on net, it is reasonable to assume that the reform was highly cost-effective.

5.2. Inference – “In-space placebo”

To get a sense of whether this pattern is something particular for Hedmark, I subject all the other regions from the donor pool to the same analysis, i.e. I construct synthetic counterfactuals in exactly the same way as above with Hedmark. Subsequently, I compute the difference between each actual region and its synthetic counterpart and plot all these difference-time series together to see whether Hedmark stands out. This plot is provided in Figure 3, from which we can see that Hedmark is clearly extreme in that for no other region do we see such a large divergence between the actual and its synthetic control taking place exactly in the treatment period.

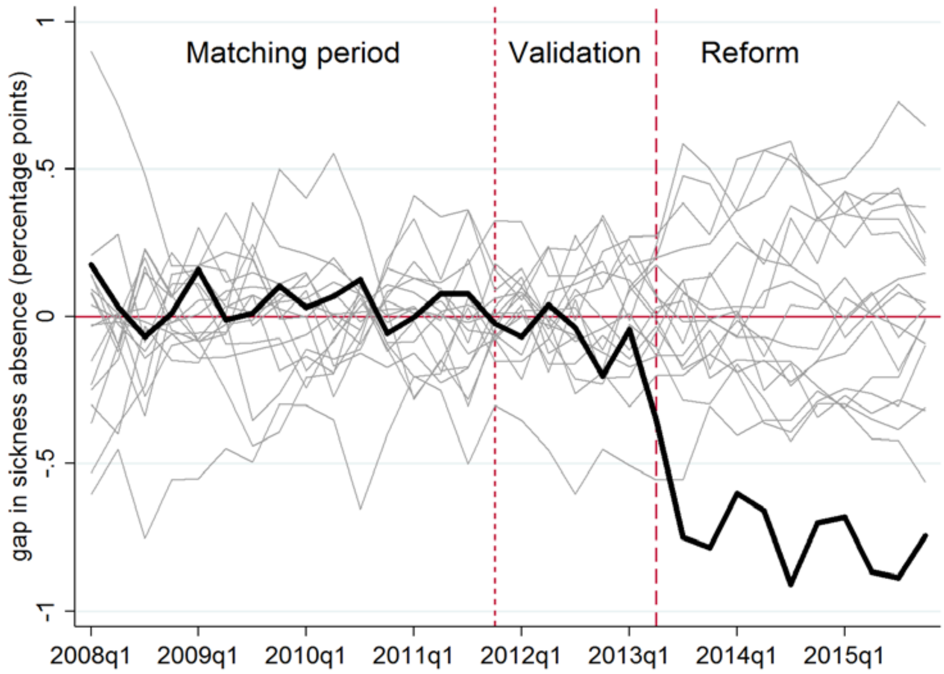

Figure 3. Difference in sickness absenteeism for all regions.
Note: The dotted line at the fourth quarter of 2012 indicates the final quarter of the matching period. The dashed line at the second quarter of 2013 indicates the period in which the activation program was introduced.

A more formal way of evaluating the preceding results for all the regions is to compare the mean squared prediction error (MSPE) from the synthetic regions for the post- versus the pre-treatment period. A relatively small MSPE indicates a good fit, while a relatively large MSPE indicates a poor fit. Comparing the MSPE before and after thus takes into how well each particular region is matched in the first place when evaluating the post-treatment behavior. A large post/pre MSPE ration thus means that the synthetic region is not well matched after relative to before treatment, or in other words that there is a break in the fit. From Figure 4 we see that Hedmark is clearly an outlier on this measure as well, with a post/pre MSPE ration of over 70, while the others range from 0-12.

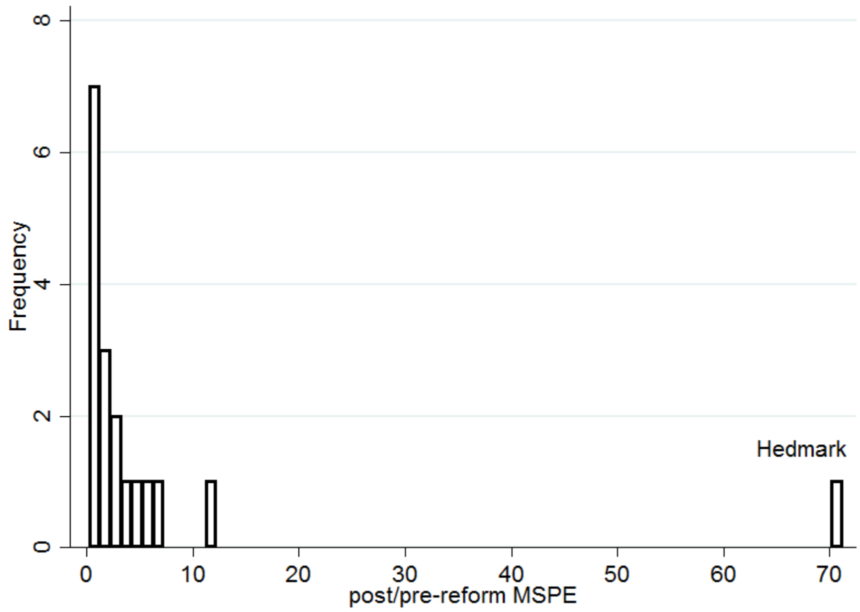

Figure 4. Distribution of ratio of mean squared prediction error post vs. pre-treatment for all regions (MSPE after treatment / MSPE before treatment).

5.3. "In-time placebo"

Following Abadie et al. (2015), I perform a so-called "in-time-placebo" study by reassigning the reform to the middle of the pretreatment period, which in this case is the third quarter of

2010, and rerunning the model. Figure 5 shows that there is no effect from the placebo reform at this point in time.

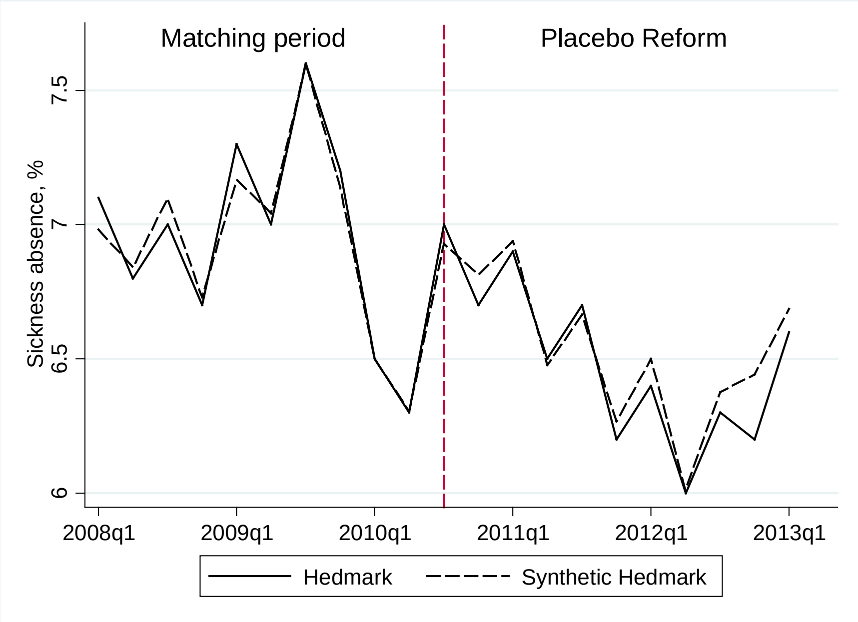

Figure 5. In-time-placebo. Estimating effect of a placebo reform taking place in the middle of the pre-treatment period.

5.4. “Leave-one-out”

Figure 6 shows that the results are robust to successively leaving out the regions receiving positive weights in the baseline analysis and performing the estimation from scratch – the resulting counterfactual post-treatment paths are very close to the baseline estimate.

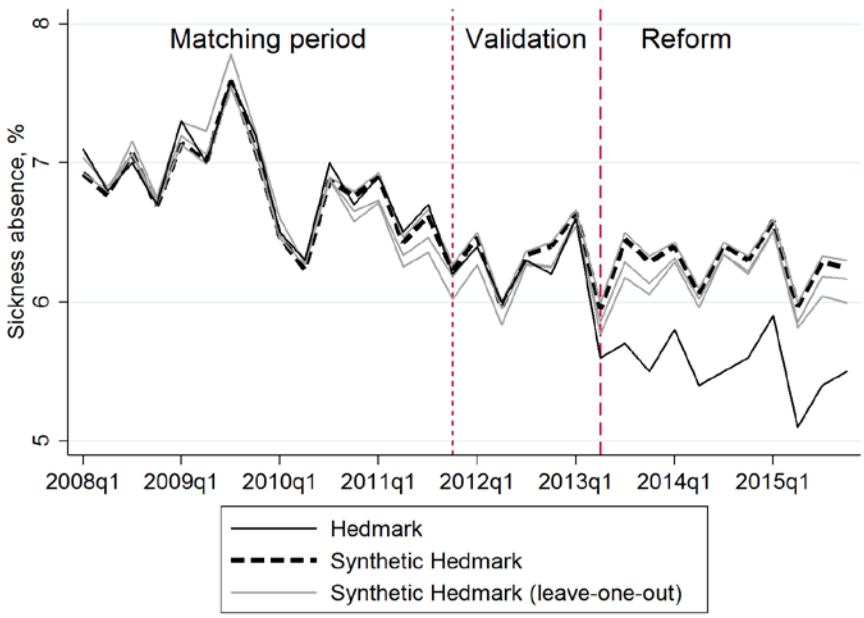

Figure 6. Leave-one-out. Excluding each of the regions with positive weight.

The results are also robust to how many and which covariates are used, as well as the length of the matching period.

5.5. Subgroup analysis

The estimated effect is similar across gender and age. Figure 7a shows the estimated effects when analyzing men and women separately. Women have a higher level of absenteeism than men, but the estimated effect is in the range of 0.5 to 1 percentage points reduction for both genders. From Figure 7b we can see that this is the case also for the three age groups of workers in their 30s, 40s, and 50s. For visibility, I here present only the difference between the actual region and its synthetic control region (along with the corresponding differences for all other regions).

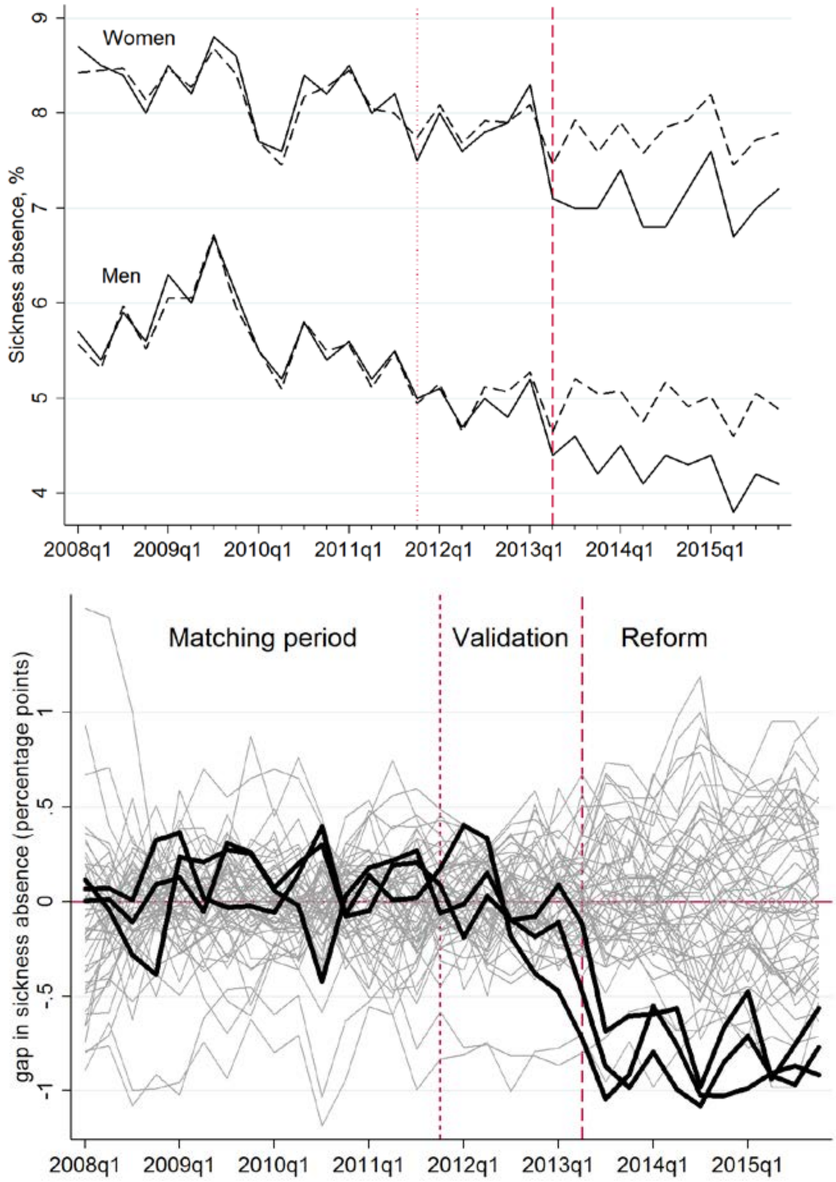

Figure 7.

a) Trends in the sickness absenteeism in Hedmark and the synthetic control region by gender (upper panel).

b) Difference in sickness absenteeism between each actual region and its synthetic control region for workers in their 30s, 40s, and 50 (lower panel).

Note: The dotted line at the fourth quarter of 2012 indicates the final quarter of the matching period. The dashed line at the second quarter of 2013 indicates the period in which the activation program was introduced.

Figure 8 presents results by industry. It is reasonable that the estimated effect is smaller for the construction sector, where many tasks are relatively more physically demanding than in other sectors, and where it may be difficult to accommodate someone not at 100% capacity. In the other sectors, the estimated reduction in absenteeism is as in the aggregate analysis up to around 1 percentage point.

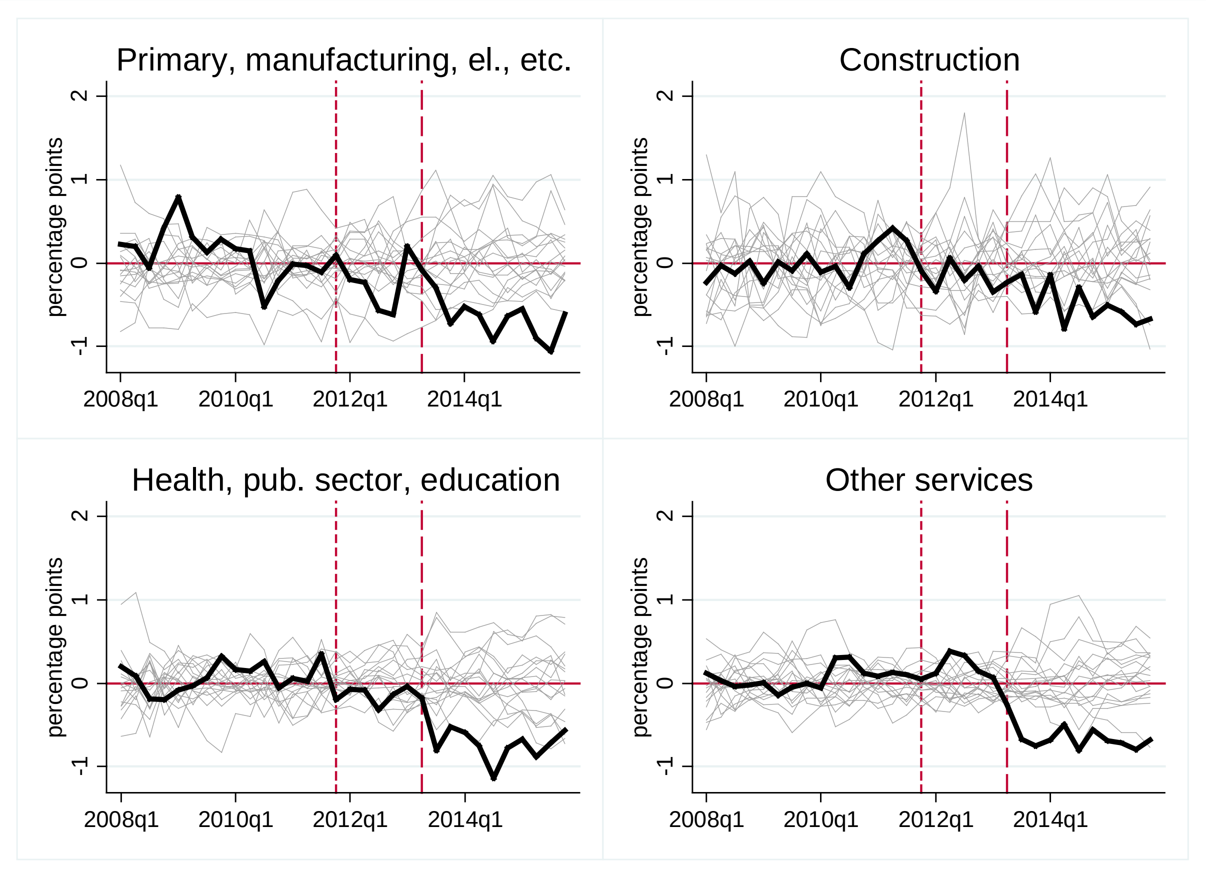

Figure 8. Difference in sickness absenteeism by industry.
Note: The dotted line at the fourth quarter of 2012 indicates the final quarter of the matching period. The dashed line at the second quarter of 2013 indicates the period in which the activation program was introduced. "Primary, etc." includes Agriculture, forestry and fishing, mining and quarrying, manufacturing, electricity, water supply, sewerage, waste management; "Health, pub. sector, education" includes public administration, defence, social security, education, human health and social work activities; "Services" includes Wholesale and retail trade, repair of motor vehicles and motorcycles, transportation and storage, accommodation and food service activities, information and communication, financial and insurance activities, real estate, professional, scientific and technical activities, administrative and support service activities, and other service activities.

6. Mechanism

Unless an exception is given, the activation requirement will typically lead to a graded sick leave certificate, meaning that the employee will be partly back at work. There is an obvious mechanical relationship between the use of graded certificates and work hours lost due to absenteeism, since a person only partly absent will still perform the rest of his hours. This mechanical effect is important in its own right, but we would also like to know whether there are other effects going on, such as speedier return to full-time work.

First, instead of taking as the dependent variable the absence rate in terms overall work hours, we can analyze the absence rate in terms of workers, i.e. the number of workers with an ongoing absence spell (regardless of grading) divided by the total number of workers. This measure is a lot less stable, since it is impacted more strongly by seasonal variation in the number of short absence spells. It can nevertheless be used to answer the question of whether the reform simply increased the presence at work of people on leave, or whether there was also a reduction in the number of people with an absence spell (full or partial). Figure 9 shows that as expected the fit is not as good as previously, however it is clear that the reform also reduced the number of claimants. The reduction is approximately 0.5 percentage points, which implies that from the base of around 80 000 employees in the region, the reform lead to around 400 fewer persons with sick pay on a given day.

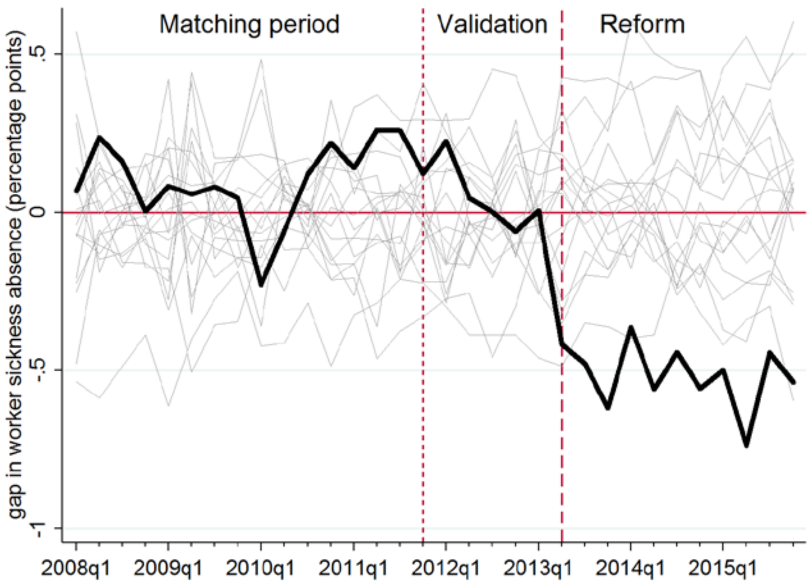

Figure 9. Difference in workers absent.
Note: The dotted line at the fourth quarter of 2012 indicates the final quarter of the matching period. The dashed line at the second quarter of 2013 indicates the period in which the activation program was introduced.

The fact that the activation requirement kicks in only after 8 weeks of absence makes it unlikely that the reform impacted the inflow to sick leave. To investigate whether accelerated termination of sick leave may have played a role, I will consider the duration of absence spells. The duration data I have available is average duration of completed spells by spell start quarter until the fourth quarter of 2013. However, since absence spells in Hedmark tended to be longer than in the rest of the country, it is hard to find a good match based on a convex combination of other regions, therefore I will here diverge somewhat from the previous analysis. From the baseline analysis of the absence rate above, I take the estimated weights presented in Table 2. Using these weights, I construct a synthetic Hedmark also for the average duration, then, knowing that the level is going to be somewhat off, subtract the values of the synthetic region from the real Hedmark to get the differences.

Figure 10 shows the results. Since also spells that started before the reform, which took place in the second quarter of 2013, may have been influenced by the new regime, a dashed line is drawn three quarters earlier (at the third quarter of 2012) to distinguish spells not affected by the reform from spells that with increasing certainty would be affected. The maximum duration is 52 weeks, so very few spells that started more than three quarters before the reform would have been influenced by it. The graph shows that if anything, there was an increasing difference in average spell duration in the period before the reform, however there is a clear break at the point at which new spells begin to be affected. This is in line with what was found by Kann et al. (2014) in their examination the duration of new spells.

Thus the reform seems to have reduced the absence rate not only through exploiting the partial work capacity of temporary disabled workers, but also by speeding up the transition rate back to full-time work.

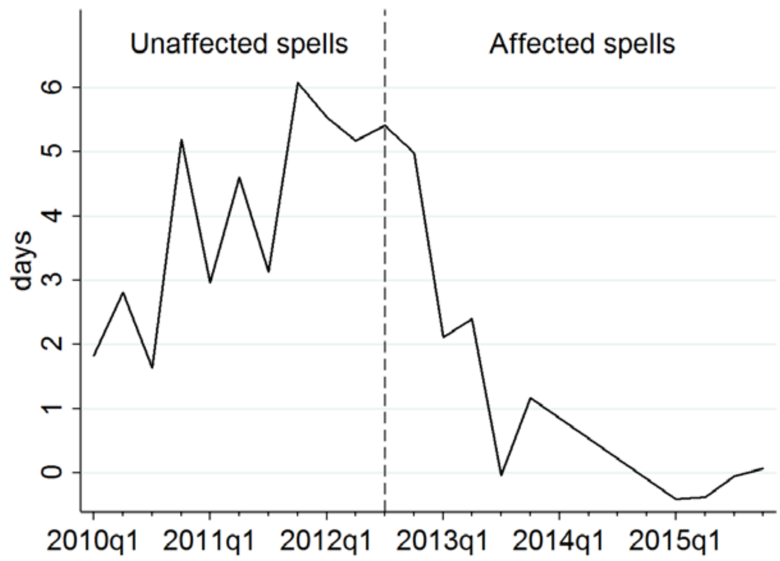

Figure 10. Difference in average spell duration by spell start quarter.

A final piece of evidence to consider is absenteeism broken down by diagnosis.[3] Musculoskeletal disorders and psychological disorders are by far the largest diagnosis groups, accounting for 40% and 20% of absenteeism, respectively. There is evidence that some degree of ac-

[3] Unfortunately, these data are only available up to and including 2014.

tivity through work is both feasible and beneficial for many types of disorders (Waddell and Burton, 2006), and in particular for musculoskeletal disorders. In contrast, for respiratory disorders, which includes infectuous diseases like the flu, presenteeism would be less desirable both from the point of view of recovery and out of concern for other workers. Thus, people suffering from respiratory diagnoses would typically not be the target of the activation requirement, and one would expect the reform to have much less of an effect on this type of absenteeism.

Figure 11 presents estimates for these three diagnosis groups and a category of other diagnoses. As expected, the largest decline occurred for absenteeism due to musculoskeletal disorders and the smallest, if any, for respiratory disorders, with diagnoses for psychological and other disorders in between.

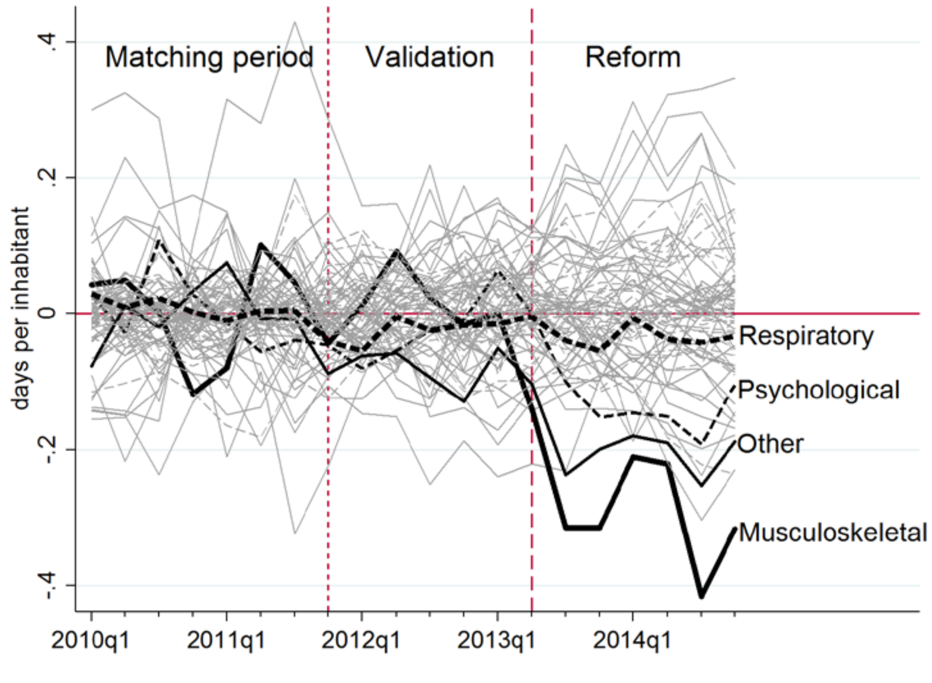

Figure 11. Difference in absenteeism by diagnosis.
Note: The dotted line at the fourth quarter of 2012 indicates the final quarter of the matching period. The dashed line at the second quarter of 2013 indicates the period in which the activation program was introduced.

7. Conclusion

In this paper, I have shown that an activation program for workers on long-term sick-leave in the Norwegian region of Hedmark reduced the level of absenteeism by 12 % compared to a synthetic control region created by a weighted average of similar regions. The results imply that exploiting their remaining work capacity of people on long-term sick leave has large potential for reducing unnecessary absenteeism and curbing costs related to social security transfers. Such an activation strategy represents an alternative to traditional attempts at welfare reform involving stricter screening or reductions in generosity, and may be more compatible with already existing legislation and contractual obligations, as well as easier to find support for across political priorities.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association* 105 (490), 493-505.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science,* 59(2), 495-510.

Abadie, A. and J. Gardeazabal (2003). The economic costs of conflict: A case study of the Basque country. *The American Economic Review* 93 (1), pp. 113-132.

Andrén, D. (2014). Does part-time sick leave help individuals with mental disorders recover lost work capacity? *Journal of Occupational Rehabilitation,* 24, 344–360.

Andrén, D., Svensson, M. (2012). Part-time sick leave as a treatment method for individuals with musculoskeletal disorders. *Journal of Occupational Rehabilitation,* 22, 418–426 .

Brage, Søren, Per Kristoffersen and Nina Lysø (2014) «NAVs oppføling av sykmeldte». *Arbeid og Velferd,* 1/2014, 78–89.

Fiva, Jon H. and Askill H. Halse (2016). Local favoritism in at-large proportional representation systems, *Journal of Public Economics*, Volume 143, November 2016, 15-26.

Hagen, E.M., Grasdal, A. and Eriksen, H.R., (2003). Does early intervention with a light mobilization program reduce long-term sick leave for low back pain: a 3-year follow-up study. *Spine* 28, 2309–2316.

Høgelund, J., Holm, A., McIntosh, J., 2010. Does graded return-to-work improve sicklisted workers' chance of returning to regular working hours? *Journal of Health Economics* 29, 158–169.

Kann, Inger Cahtrine, Ivar Andreas Åsland Lima and Per Kristoffersen (2014). Håndheving av aktivitetskravet I Hedmark har redusert sykefraværet. *Arbeid og velferd* (3), pp. 14-33.

Kausto, J., Solovieva, S., Virta, L. J., and Viikari-Juntura, E. (2012). Partial sick leave associated with disability pension: propensity score approach in a register-based cohort study. *BMJ Open*, 2(6), e001752. http://doi.org/10.1136/bmjopen-2012-001752.

Kausto, J., Viikari-Juntura, E., Virta, L. J., Gould, R., Koskinen, A., & Solovieva, S. (2014). Effectiveness of new legislation on partial sickness benefit on work participation: a quasi-experiment in Finland. *BMJ Open*, 4(12), e006685. http://doi.org/10.1136/bmjopen-2014-006685.

Markussen, S. (2012). The individual cost of sick leave. *Journal of Population Economics*, 25(4), 1287-1306.

Markussen, S., Mykletun, A., & Røed, K. (2012). The case for presenteeism—Evidence from Norway's sickness insurance program. *Journal of Public Economics*, 96(11), 959-972.

Mykletun, A., Eriksen, H.R., Røed, K., Schmidt, G., Fosse, A., Damberg, G., Christiansen, E.C., Guldvog, B., (2010). Tiltak for reduksjon i sykefravær: Aktiviserings - og nærværsreform. Ekspertgrupperapport til Arbeidsdepartementet. English Summary available at: http://www.regjeringen.no/en/dep/aid/topics/welfare-policy/inclusive-working-life/summary-of-.html?id=592754.

NAV Hedmark (2014) Sykefraværssatsingen i NAV Hedmark. Arbeid knyttet til aktivitetskravet for sykmeldte. *Intern rapport, NAV Hedmark. [Internal report, NAV Hedmark.]*

OECD (2008). Are all jobs good for your health? The impact of work status and working conditions on mental health. Chapter 4 in OECD Employment Outlook 2008 . Paris.

OECD (2010). Sickness, Disability and Work: Breaking the Barriers — Key Trends and Outcomes in Sickness and Disability. OECD, Paris

Røed, K. (2012). Active social insurance. *IZA Journal of Labor Policy*, 1(1), 1.

Van Vilsteren, M., Van Oostrom, S.H., de Vet, H.C.W., Franche, R.L., Boot C.R., Anema, J.R. (2015). Workplace Interventions to Prevent Work Disability in Workers on Sick Leave (Review). *Cochrane Database of Systematic Reviews* 2015, Issue 10.

Viikari-Juntura, E., Kausto, J., Shiri, R., Kaila-Kangas, L., Takala, E.-P., Karppinen, J., Miranda, H., Luukkonen, R., Martimo, K.-P. (2012). Return to work after early parttime sick leave due to musculoskeletal disorders: a randomized controlled trial. *Scandinavian Journal of Work, Environment & Health* 38 (2), 134–143.

Waddell, G., Burton, A.K., (2006). Is work good for your health and well-being? Report for Department for Work and Pensions. TSO (The Stationery Office), London.

Appendix

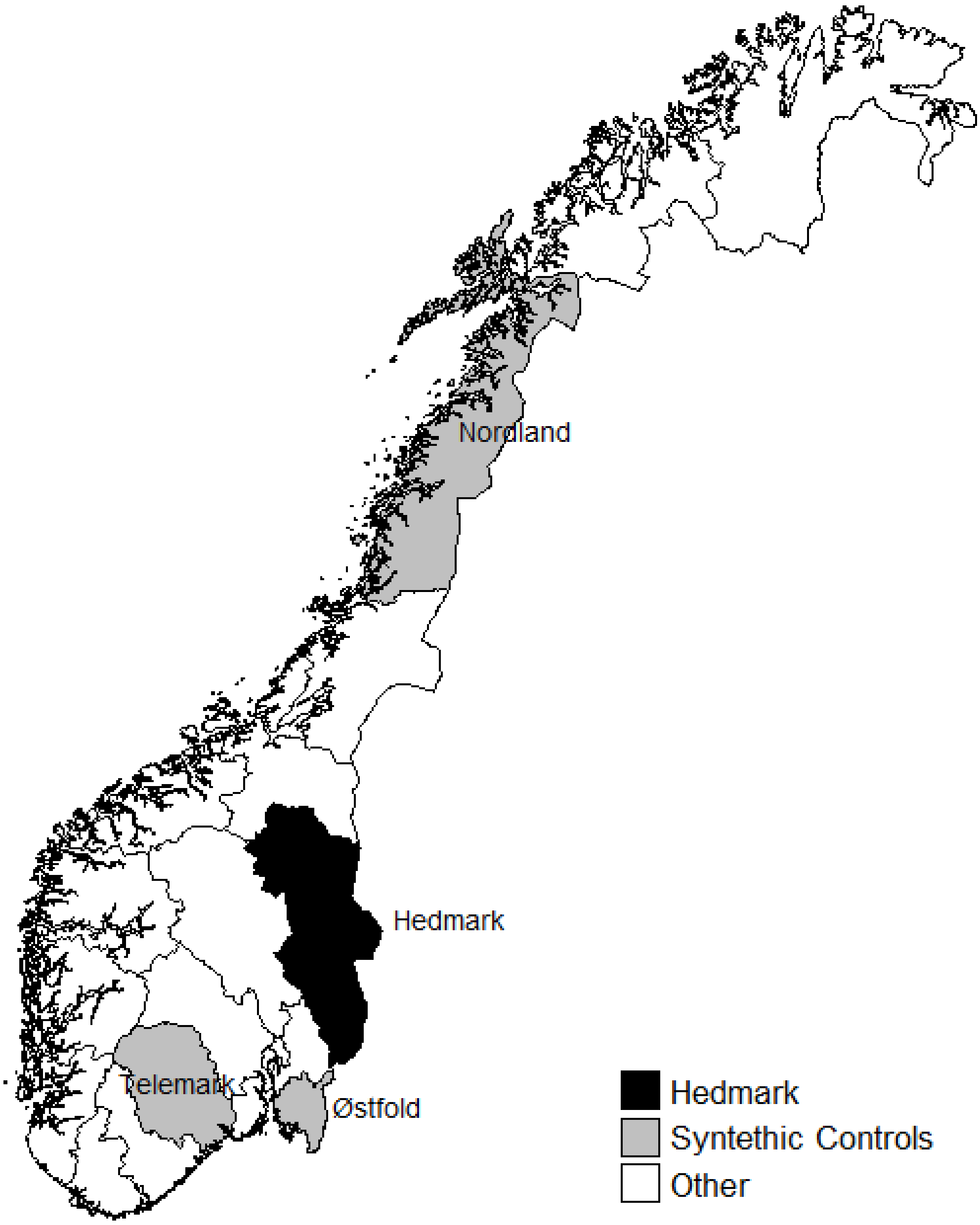

Figure A.1. Norwegian regions.